American Journal of Epidemiology

Vol. 173, No. 9
DOI: 10.1093/aje/kwq481
Advance Access publication:
March 21, 2011

Practice of Epidemiology

Bias Due to Left Truncation and Left Censoring in Longitudinal Studies of Developmental and Disease Processes

Kevin C. Cain*, Siobán D. Harlow, Roderick J. Little, Bin Nan, Matheos Yosef, John R. Taffe, and Michael R. Elliott

* Correspondence to Dr. Kevin C. Cain, Department of Biostatistics, University of Washington, Box 357232, F-600 Health Sciences Building, Seattle, WA 98195-7232 (e-mail: cain@uw.edu).

Initially submitted June 8, 2009; accepted for publication December 8, 2010.

In longitudinal studies of developmental and disease processes, participants are followed prospectively with intermediate milestones identified as they occur. Frequently, studies enroll participants over a range of ages including ages at which some participants' milestones have already passed. Ages at milestones that occur prior to study entry are left censored if individuals are enrolled in the study or left truncated if they are not. The authors examined the bias incurred by ignoring these issues when estimating the distribution of age at milestones or the time between 2 milestones. Methods that account for left truncation and censoring are considered. Data on the menopausal transition are used to illustrate the problem. Simulations show that bias can be substantial and that standard errors can be severely underestimated in naïve analyses that ignore left truncation. Bias can be reduced when analyses account for left truncation, although the results are unstable when the fraction truncated is high. Simulations suggest that a better solution, when possible, is to modify the study design so that information on current status (i.e., whether or not a milestone has passed) is collected on all potential participants, analyzing those who are past the milestone at the time of recruitment as left censored rather than excluding such individuals from the analysis.

bias (epidemiology); censoring; epidemiologic methods; longitudinal studies; study design; truncation

Studies of developmental and disease processes often follow participants prospectively, identifying milestones as they occur. Examples include longitudinal studies of the developmental stages in puberty, timing of pregnancies, time to spontaneous abortion, presentation of occupational disease in workers, development of sequelae of chronic diseases, and stages of reproductive aging among midlife women. Ideally, all study participants would be enrolled prior to the first milestone of interest and followed until the final milestone was observed, so the entire process is observed for all participants. However, this so-called incident cohort design (1) is not always feasible. The process of interest may develop over many years, and the age at onset of the process or timing of milestones may vary considerably across participants.

In practice, researchers usually recruit participants over a range of ages when most are expected to be near the first milestone of interest. Individuals who have already passed the final milestone or a specified earlier milestone at the time of recruitment are usually excluded from the study. Such a design, referred to as a prevalent cohort study (1), is subject to bias from 3 causes, if not analyzed correctly (2). Right censoring occurs when a participant has not yet reached the milestone of interest at study end. Left censoring occurs if a participant is entered into the study when the milestone of interest occurred prior to study entry but the age at that milestone is unknown. Left truncation occurs when individuals who have already passed the milestone at the time of study recruitment are not included in the study. Survival analysis methods for dealing with right censoring (2, 3) are widely understood and implemented by epidemiologists. However, the methods for adjusting for left truncation and left censoring are less widely known and infrequently applied in longitudinal epidemiologic studies.

When a single milestone is subject to truncation, it is fairly easy to understand why truncation causes bias. For example, recent articles (4, 5) have discussed the large bias caused by left truncation in estimating the risk of spontaneous abortion.

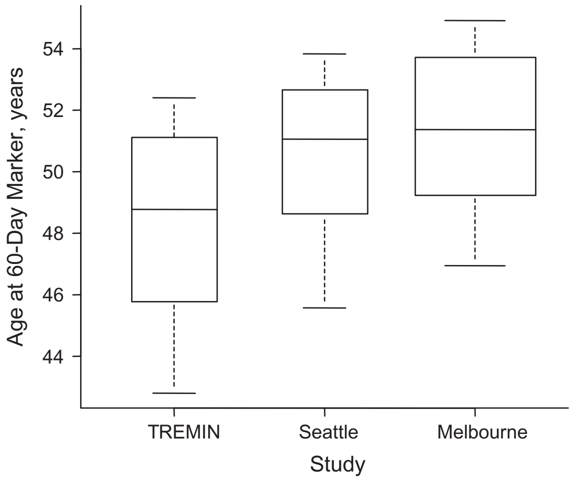

Figure 1. Boxplots of the estimated distribution of age at onset of the late menopausal transition, as defined by first occurrence of a menstrual cycle of at least 60 days in length. Data from 3 studies are shown. Modified from a figure in Harlow et al. (6).

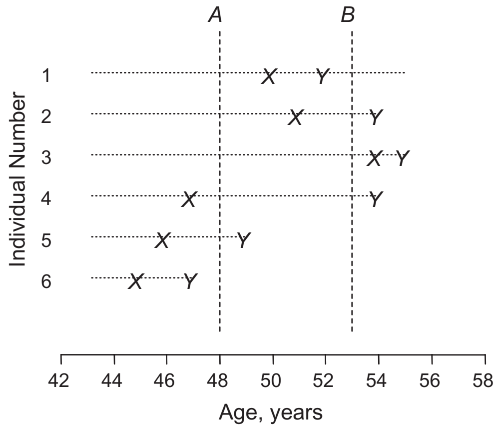

Figure 2. Timing of milestones X and Y relative to window $A - B$. In this hypothetical situation, individuals are recruited into the study at age 48 years (A) and followed until age 53 years (B). Six individuals are portrayed. In reality, A and B differ by participant but, for simplicity of presentation, the figure shows A and B as being the same for all 6 individuals. The individual in row 1 had milestone X at age 50 years and milestone Y at age 52 years. Both events are within the observation window and, hence, both milestones are observed. For the second individual, X is observed, but Y has not yet occurred when observation ends at B, while for the third, both X and Y occur after observation ends. For the 3 remaining individuals, X occurred before study entry, and for the sixth person Y also occurred before study entry.

In this setting, a large fraction of pregnant women enter the study after the gestational age corresponding to peak risk of spontaneous abortion, leading to an underestimation of risk if the analysis does not account for the fact that many women with an early spontaneous abortion have been excluded. Spontaneous abortions often occur before women know they are pregnant and are undetected.

A more complex situation involving 2 milestones is illustrated by a study evaluating different methods for defining onset of the late stage of the menopausal transition in midlife women. Figure 1, modified from a figure in Harlow et al. (6), shows the estimated distribution of age at onset of the late menopausal transition in 3 studies, as defined by the occurrence of a menstrual cycle of at least 60 days in length after age 40 years. A striking feature is that the estimated ages at late transition are much younger for the TREMIN Research Program on Women's Health (''TREMIN'') than for the Melbourne Women's Midlife Health Project (''Melbourne'') study, with estimated ages for the Seattle Midlife Women's Health (''Seattle'') study falling in between. TREMIN (7) participants all enrolled prior to age 35 years, while the age ranges for study entry were 35–55 for Seattle (8) and 45–55 for Melbourne (9). In the latter 2 studies, women who were already postmenopause at the time of recruitment were excluded. Thus, the higher estimates of age at onset of the menopausal transition in these studies could arise because the estimation procedure fails to account for the following: 1) left censoring of age at late transition because for some women the late transition had occurred prior to study entry; and 2) left truncation on age at final menstrual period because women already past the final menstrual period are excluded.

Information on the excluded cases could potentially be collected and used in the analyses. We focus mainly on the situation where such information is absent, but we also study the potential gain from collecting information on the excluded cases. We consider the single-milestone case but also address the more complicated case where truncation is based on 1 milestone, but the variable of interest is either the time of a different milestone or the time between 2 milestones. The menopausal transition is used throughout as the motivating example, with 2 milestones: the transition from early to late stage (6, 10, 11) of the menopausal transition and the final menstrual period. Simulations to estimate bias are based on longitudinal menstrual data from TREMIN (7). Analyses that account for left truncation and/or left censoring are presented, as well as naïve analyses that ignore these features.

MATERIALS AND METHODS

Let Y denote age at a terminal event or milestone, such as the final menstrual period. Let X denote age at an intermediate event or milestone, such as entry to the late stage of the menopausal transition. By definition, $X \leq Y$. Let D denote the difference $(Y - X)$, representing years spent in late menopausal transition in our example. Let A denote age at potential study entry and B denote age at potential study exit. The goals of analysis are to estimate the distributions of Y, X, and D.

Figure 2 illustrates the possible timings of milestones X and Y relative to the ages at entry (A) and exit (B) from the study. Which of the 6 types of individuals will be included in the study depends on the inclusion criteria specified in the study design. In this paper, we consider 2 common study designs. In Design I, people who have already passed milestone X, like participant types 4–6 in Figure 2, are excluded, which means that truncation acts on variable X. In the example of transition to menopause, this means that women who are already in late transition at A are excluded.

In Design II, participants who have already passed milestone Y, like participant type 6, are excluded. Participants like 4 and 5 are included, but the value of X is left censored. Thus, milestone X is subject to left censoring, while truncation acts on milestone Y. In our example, postmenopausal women are excluded under Design II, and women already in late stage at the time of entry are included but left censored.

The single-milestone case—estimating the distribution of *X* in Design I

Under Design I, left truncation acts on X: Subjects with $X < A$ are excluded. The researcher knows that such individuals would have been excluded but does not know what number are excluded. If A is the same for all participants, then an estimator that ignores left truncation will estimate the distribution of X conditional on $X > A$, which differs from the unconditional or marginal distribution of X. When age at study entry A is variable, a version of the nonparametric product-limit estimator that incorporates age at entry will correctly deal with left truncation to give an unbiased nonparametric estimate of the distribution of X. Code for doing such a nonparametric analysis in STATA (StataCorp LP, College Station, Texas), SAS (SAS Institute, Inc., Cary, North Carolina), and R (R Foundation for Statistical Computing, Vienna, Austria) statistical software/language is provided in Web Appendix 1, the first of 4 Web appendices posted on the *Journal*'s Web site (http://aje.oxfordjournals.org/). This procedure gives an unbiased estimate of the unconditional distribution of X only if the lowest A in the data set (A_{min}) is earlier than the earliest possible value of X. Moreover, the estimate can be very unstable if the first X is observed shortly after A_{min}, when the risk set is small because only a few participants have entered prior to that age. Precision can be increased with a modification to estimate the distribution of X conditional on $X > A_0$, where A_0 is some value greater than A_{min}, chosen so that a fairly large number of participants have $A < A_0$. However, the improvement in precision comes at the cost of increased bias, to the extent that the conditional and unconditional distributions of X differ.

When a large fraction of the left tail of the distribution of X is below the lowest age at entry, A_{min} (or A_0), then the only way to get unbiased estimates of the unconditional distribution is to know the appropriate parametric distribution for X, so that the probability in that truncated left tail can be estimated. If a parametric distribution is posited, it is fairly easy to modify the likelihood to account for left truncation: The likelihood is divided by the probability that the observation is included in the sample given the age at entry for each participant, i.e., $\Pr(X > A)$. Web Appendix 1 provides STATA and R code for this analysis, assuming a normal distribution.

Now consider an alternative to Design I, in which current status information (i.e., whether or not X has occurred prior to A) is recorded on all potential participants. In this situation, participants with $X < A$ can be included in the analysis as left censored. Web Appendix 1 provides code for analysis of left-censored data.

Both of the designs above assume that one can reliably determine whether or not X has occurred prior to study entry. Unfortunately, this is not always the case. In the example of the menopausal transition, one definition of onset of late transition is the first occurrence of a long (e.g., 60 days or more) menstrual cycle, defined as days from the start of one menstrual episode to the start of the next episode, after age 40 years. If a woman enters the study at age 48 years and a long menstrual cycle occurs some time later, it may not be clear whether this was the first occurrence of a long cycle or whether one or more long cycles actually occurred prior to study entry. It is tempting, though unwise, to just treat the first observed occurrence as if it were the actual first occurrence, as was done in the latter 2 studies portrayed in Figure 1 (and which we refer to as the naïve method in the simulation results for Design II below).

The double-milestone case—estimation of the distributions of *X* and *D* under Design II

The more complicated 2-milestone case involves variables X, Y, and $D = Y - X$. Under Design II, truncation acts on Y, and X is subject to left censoring. Analyses of X and D are potentially influenced by both of these effects. An appropriate analysis requires a model for the joint distribution of X and Y. Truncation based on Y will cause bias in the estimated distribution of X only if Y and X are not independent. Jiang et al. (12) discuss truncation and censoring in the context of the development of diabetic retinopathy (X) and death (Y) in patients with diabetes. They propose a semiparametric model in which the association between X and Y is described by a parametric relation, while the marginal distributions of X and Y are estimated nonparametrically.

We propose a different model that postulates a bivariate normal distribution for X and Y. We have developed R functions (provided in Web Appendix 2) to find the maximum likelihood estimates for the means of X, Y, and D while accounting for either the left truncation of X (Model I) or left censoring of X and left truncation of Y (Model II). A fully parametric model makes stronger distributional assumptions but facilitates estimation of the distribution of D as well as the distributions of X and Y.

Simulations based on menopausal transition data

The TREMIN study (7) enrolled 1,997 women students at the University of Minnesota in 1935–1939, many of whom maintained menstrual calendars throughout their reproductive life. Our analysis includes data from age 40 years onward for the 660 women who were still participating at age 40. The menstrual calendars provided the length in days of their menstrual cycles. $Y =$ age at the final menstrual period was defined as the age at the last bleeding episode preceding 12 bleed-free months. $X =$ age at the start of the late menopausal transition was defined as the first observed cycle of length 60 days or longer, with the further condition that isolated 60-day cycles (with no recurrences in the next 10 cycles) before age 45 years do not count as the start of late stage (13). Four women appear to be outliers, showing persistent long cycles prior to age 45 years but the final menstrual period much later. After removal of these outliers, an examination of the marginal distributions of X and Y indicates nearly normal distributions (Web Appendix 3). However, a deviation from bivariate normality is apparent with D having a skewed distribution, in part because of the fact that no points have $X > Y$, by definition. Nonetheless, we proceed with analyses based on the bivariate normal assumption and discuss in Web Appendix 3 the possible methods for improving the model.

A useful feature of TREMIN data is the absence of left censoring or left truncation. Thus, we can estimate the distributions of X, Y, and $D = Y - X$ on the basis of the full data set with no truncation and then assess bias by simulating various amounts of truncation. The simulated data sets are generated as follows:

1. A 10-year range for age at entry is specified, for example, 40–50 years.
2. A bootstrap (14) sample is drawn from the original data set.
3. For each participant, a random number A is generated from a uniform distribution over the 10-year age range.

4a. For Design I, if $X < A$, then that participant is deleted from the data set.

4b. For Design II, if $Y < A$, then that participant is deleted from the data set.

5. The reduced data set is used to produce estimates of the mean and standard deviation of X and D based on several alternative methods.

Steps 2–5 are repeated 100 times for each age range, and then the process is repeated for different age ranges at entry from 35–45 to 47.5–57.5 years. Simulations and data analyses use R, version 2.10.1, software.

RESULTS

Scatterplots of simulation results are displayed in Web Figures 11–16 in Web Appendix 4, for both parametric and nonparametric estimators. Simulation results for parametric estimators of the mean are summarized in the following tables; patterns of results are similar for nonparametric estimators of the median.

Table 1 shows maximum likelihood estimates of the mean of X, age at entry to the late stage of the menopausal transition, based on simulations of Design I that has truncation on X; refer to Web Figure 12 in Web Appendix 4 for a graphical display of these results. Rows in Table 1 show averages of the mean and averages of the standard error over 100 simulated data sets. The first row of results is based on the full data set with no truncation, giving an estimated mean age at the start of late stage of 48.99 years. Subsequent rows show how the estimates of the mean and standard error change as the amount of truncation increases (i.e., age at entry increases), as well as the empirical bias in years, defined as the mean age at late transition minus the value in the absence of truncation (row 1). Results of 3 analyses are shown in Table 1. All analyses are the normal-based maximum likelihood estimate and account for right censoring. Analysis 1 ignores left truncation. The likelihood for Analysis 2 accounts for left truncation by dividing by $\Pr(X > A)$. Analysis 3 is based on the hypothetical alternative study design in which observations with X prior to study entry are included in the analysis as left censored, rather than being left truncated and excluded.

As the amount of truncation increases, the naïve estimate has an increasing positive empirical bias. For age at entry of 42.5–52.5 years, the empirical bias for the naïve estimator is substantial, namely, 1.29 years. Contrast this bias with the empirical bias of only 0.07 for Analysis 2, which controls for left truncation, and −0.01 for Analysis 3, which treats observations as left censored rather than left truncated. Approximately half of the participants have been truncated for

Table 1. Estimates of Age at Entry to Late Stage Based on Design I, Left Truncated on X[a]

Simulated Age at Study Entry, years		No. of Participants	No. of Participants Observed[b]	Analysis 1 (Naïve, Ignores Left Truncation on X)		Analysis 2 (Accounts for Left Truncation on X)		Analysis 3[c] (Accounts for Left Censoring of X)	
Median	Range			Mean (SE)	Empirical Bias[d]	Mean (SE)	Empirical Bias	Mean (SE)	Empirical Bias
	No truncation	656	462	48.99 (0.14)	0.00	48.99 (0.14)	0.00	48.99 (0.14)	0.00
40.0	35–45	623	452	49.05 (0.14)	0.06	48.96 (0.15)	−0.03	48.95 (0.14)	−0.04
42.5	37.5–47.5	565	424	49.31 (0.14)	0.32	49.00 (0.16)	0.01	48.97 (0.14)	−0.02
45.0	40–50	465	362	49.73 (0.14)	0.74	49.01 (0.19)	0.02	48.98 (0.15)	−0.01
47.5	42.5–52.5	334	269	50.28 (0.16)	1.29	49.06 (0.26)	0.07	48.98 (0.16)	−0.01
50.0	45–55	199	166	50.97 (0.19)	1.98	48.37 (0.77)	−0.62	48.94 (0.20)	−0.05
52.5	47.5–57.5	96	80	52.07 (0.24)	3.08	37.82 (6.67)	−11.17	48.94 (0.30)	−0.05

Abbreviations: SE, standard error; X, age at entry to late stage, first 60-day cycle.

[a] All analyses account for right censoring.

[b] Number of participants with age at entry to late stage observed rather than being right censored.

[c] Analysis 3 is based on the hypothetical alternative study design in which observations with X prior to study entry are included in the study analysis as left censored, rather than being left truncated and excluded from analyses as in Analyses 1 and 2.

[d] Defined as the mean on that row minus the mean on the first row (which has no truncation).

this row and, thus, it is not surprising that the naïve methods have substantial bias. Note also the differences in standard error. Standard errors increase rapidly for Analysis 2, as the age at simulated study entry increases, appropriately reflecting the fact that the available information is greatly reduced. In contrast, the standard error for the naïve parametric estimators increases much more slowly: The precision of the estimated mean is overstated. The standard error for Analysis 3 increases less slowly than for Analysis 2, reflecting the fact that knowledge of left censored observations adds information to the analysis.

Table 2 shows estimates for age at onset of late transition (X) for Design II with left truncation on age at the final menstrual period (Y) and X subject to left censoring (Web Figure 15 in Web Appendix 4). All estimates are based on a bivariate normal distribution of X and Y. The naïve estimates ignore truncation and do not deal correctly with left censoring on X. Specifically, the naïve approach sets X equal to age at the first observed 60-day cycle after A, corresponding to what would be done if the first observed 60-day cycle after study entry were assumed to be the first in that woman's lifetime. In the row corresponding to median simulated age at entry of 47.5 years, the naïve estimator has an empirical bias of 1.13, slightly less than under Design I (Table 1).

Accounting for left censoring on X in Analysis 2 reduces bias by about one half. In addition, accounting for left truncation on Y in Analysis 3 reduces bias almost to zero, except for the final row in which the severe truncation leads to unstable results and a high standard error. The final column of Table 2 shows results based on an alternative design in which observations are left censored on Y rather than being left truncated on Y. Compared with Analysis 3, the standard error is smaller under this alternative design, and bias is small except for the final 2 rows.

Table 3 (also Web Figure 16 in Web Appendix 4) shows estimates of the time spent in late stage ($D = Y - X$) based on Design II and a bivariate normal distribution for X and Y. The empirical bias of the naïve Analysis 1 estimates is increasingly negative as the degree of truncation increases. However, the bias for Analysis 2, which accounts for left censoring of X, is positive but rather small. This latter result indicates that the large negative bias in Analysis 1 is primarily due to failure to account for left censoring of X.

The remaining small positive bias of Analysis 2 is due to truncation on Y and the fact that Y and D are positively correlated. This bias is reduced almost to zero (except in the bottom row) for both Analysis 3, which accounts for truncation, and Analysis 4, which considers observations as left censored on Y rather than truncated.

DISCUSSION

Left truncation in studies of developmental processes is not just of theoretical interest: It can cause substantial bias if ignored. Examples in which a large fraction of potential observations are left truncated are rate of spontaneous abortion (4) and age at menopause transition stages (6). The difference in Figure 1 between the Melbourne study, with

Table 2. Estimates of Age at Entry to Late Stage Based on Design II, Left Truncated on Y and Left Censored on X[a]

Simulated Age at Study Entry, years		No. of Participants	No. of Participants Observed[b]	Analysis 1 (Naïve, Ignores Truncation on Y and Left Censoring of X)		Analysis 2 (Accounts for Left Censoring of X, Ignores Left Truncation on Y)		Analysis 3 (Accounts for Left Censoring of X and Left Truncation on Y)		Analysis 4[c] (Accounts for Left Censoring of X and Left Censoring of Y)	
Median	Range			Mean (SE)	Empirical Bias[d]	Mean (SE)	Empirical Bias	Mean (SE)	Empirical Bias	Mean (SE)	Empirical Bias
	No truncation	656	462	48.99 (0.14)	0.00	48.99 (0.14)	0.00	48.99 (0.14)	0.00	48.99 (0.14)	0.00
40	35–45	631	458	49.00 (0.14)	0.01	48.97 (0.14)	−0.02	48.96 (0.14)	−0.03	48.95 (0.14)	−0.04
42.5	37.5–47.5	593	450	49.17 (0.13)	0.18	49.07 (0.14)	0.08	48.98 (0.15)	−0.01	48.97 (0.14)	−0.02
45	40–50	525	417	49.54 (0.13)	0.55	49.26 (0.14)	0.27	49.00 (0.16)	0.01	48.98 (0.15)	−0.01
47.5	42.5–52.5	421	348	50.12 (0.14)	1.13	49.58 (0.15)	0.59	49.03 (0.19)	0.04	48.96 (0.16)	−0.03
50	45–55	294	252	50.83 (0.15)	1.84	49.99 (0.18)	1.00	49.02 (0.28)	0.03	48.86 (0.19)	−0.13
52.5	47.5–57.5	174	150	51.83 (0.18)	2.84	50.65 (0.23)	1.66	48.83 (0.64)	−0.16	48.85 (0.26)	−0.14

Abbreviations: SE, standard error; X, age at entry to late stage, first 60-day cycle; Y, age at final menstrual period.

[a] All analyses account for right censoring.

[b] Number of participants with age at entry to late stage observed rather than being right censored.

[c] Analysis 4 is based on the hypothetical alternative study design in which observations with Y prior to study entry are included in the study analysis as left censored, rather than being left truncated and excluded from analyses as in Analyses 1–3.

[d] Defined as the mean on that row minus the mean on the first row (which has no truncation).

Table 3. Estimates of Years in Late Stage ($D = Y - X$) Based on Design II, Left Truncation on Y and Left Censoring of X[a]

Simulated Age at Study Entry, years		No. of Participants	No. of Participants Observed[b]	Analysis 1 (Naïve, Ignores Truncation on Y and Left Censoring of X)		Analysis 2 (Accounts for Left Censoring of X, Ignores Left Truncation on Y)		Analysis 3 (Accounts for Left Censoring of X and Left Truncation on Y)		Analysis 4[c] (Accounts for Left Censoring of X and Left Censoring of Y)	
Median	Range			Mean (SE)	Empirical Bias[d]	Mean (SE)	Empirical Bias	Mean (SE)	Empirical Bias	Mean (SE)	Empirical Bias
	No truncation	656	462	2.75 (0.09)	0.00	2.75 (0.09)	0.00	2.75 (0.09)	0.00	2.75 (0.09)	0.00
40	35–45	631	325	2.72 (0.09)	−0.03	2.74 (0.09)	−0.01	2.74 (0.09)	−0.01	2.74 (0.09)	−0.01
42.5	37.5–47.5	593	319	2.66 (0.09)	−0.09	2.77 (0.09)	0.02	2.75 (0.09)	0.00	2.75 (0.09)	0.00
45	40–50	525	295	2.49 (0.09)	−0.26	2.78 (0.09)	0.03	2.73 (0.09)	−0.02	2.74 (0.09)	−0.01
47.5	42.5–52.5	421	249	2.29 (0.09)	−0.46	2.84 (0.10)	0.09	2.73 (0.10)	−0.02	2.73 (0.10)	−0.02
50	45–55	294	178	2.06 (0.10)	−0.69	2.93 (0.13)	0.18	2.71 (0.14)	−0.04	2.70 (0.13)	−0.05
52.5	47.5–57.5	174	106	1.79 (0.12)	−0.96	3.01 (0.18)	0.26	2.61 (0.25)	−0.14	2.63 (0.19)	−0.12

Abbreviations: D, years in late stage; SE, standard error; X, age at marker for late stage, first 60-day cycle; Y, age at final menstrual period.

[a] All analyses account for right censoring.

[b] Number of participants with years in late stage observed rather than right censored; that is, both age at entry to late stage and age at final menstrual period are observed.

[c] Analysis 4 is based on the hypothetical alternative study design in which observations with Y prior to study entry are included in the study analysis as left censored, rather than being left truncated and excluded from analyses as in Analyses 1–3.

[d] Defined as the mean on that row minus the mean on the first row (which has no truncation).

age at entry of 45–55 years, and the TREMIN study is a little over 2 years. The biases of the naïve analyses simulated in Tables 1 and 2 for the row corresponding to age at entry of 45–55 years are 1.98 and 1.84 years, respectively, indicating that the difference seen in Figure 1 could be attributable mostly to failure to correctly account for left truncation on age at the final menstrual period and left censoring of age at onset of late transition.

Analyses to deal with truncation are fairly straightforward when the variable of interest is also the variable that determines truncation. Both nonparametric and parametric methods exist in readily available software (R, SAS, STATA) for dealing with truncation in such cases. However, when the distribution of 1 milestone (X) is of interest but truncation acts on a second milestone (Y), it is necessary to model the association between them, and such a model requires at least some parametric assumptions. Parametric models based on a normal or other distribution can be useful in this situation, although results are dependent on the distributional and model assumptions. The bivariate normal model used for our example works quite well for estimating the distribution of X, which is very close to a normal distribution.

The quantity D, defined as $Y - X$, has a nonnormal distribution in the TREMIN data. The simulations presented here are useful for illustrating the existence of bias and how it can be reduced by accounting for truncation and left censoring, but alternative models that modify the normal assumption are desirable if one is interested in accurately estimating the distribution of D. Web Appendix 3 contains an example of such an exploration and a discussion of some challenges.

It should be noted that a limitation of any distribution-based analysis is that results could be sensitive to distributional assumptions, yet checking these assumptions may be impossible if a large fraction of potential participants have been truncated such that the leftmost part the distribution is not represented in the sample. Our example is atypical in that untruncated data are available, so it is possible to evaluate the distributional assumptions used in the simulations.

Our simulations show that accounting for left truncation greatly reduces bias, but estimates become unstable as the amount of truncation approaches or exceeds 50%. This instability is to be expected. In the extreme case, one is estimating the entire distribution on the basis of data from only the right tail of the distribution, which is probably unwise. Another possible concern about analyses that account for left truncation is that they assume the truncation boundary is sharp. For truncation on Y, this means that, for individuals with $Y \leq A$, the probability of being included is zero, and for those with $Y > A$ the probability of being included is unrelated to how large $(Y - A)$ is, for example, how close a woman is to menopause. This assumption may not be true in some settings. Further investigation is needed into the effect of a nonsharp inclusion boundary on estimates that account for truncation, and whether the truncation adjustment can be modified to account for a nonsharp boundary.

In some situations, it is possible to address the problem of truncation by modifying the study design. Individuals with $Y < A$ are not always identifiable (e.g., in spontaneous abortion), but when they are identifiable it is preferable to collect current status information (i.e., whether Y has occurred yet) on all potential study participants at initial assessment. Combining these cross-sectional data with longitudinal data on participants with $Y > A$ converts a left-truncation problem into a left-censoring problem and leads to estimates that are less biased, more precise, and more robust to high degrees of censoring. Knowing the number of observations with $Y < A$ is more informative than just knowing that left truncation exists but with no knowledge of how many individuals have been truncated.

This paper has focused on estimating the marginal or unconditional distributions of X, Y, and D. Other types of analysis are also of interest, for example, the association of predictors, such as body mass index and smoking status, with X, D, or Y. Web Appendix 1 gives code for doing such an analysis, while accounting for left truncation or left censoring in the single-milestone case.

The bias due to truncation and censoring discussed in this report is an example of the more general problem that structural biases can occur when the inclusion criteria for a study are related either directly or indirectly to a variable of primary study interest. Many other examples of this phenomenon exist, such as cancer screening and longitudinal studies of occupational exposures. In any such setting, analyses should consider the study-specific details of recruitment and inclusion criteria that can lead to bias. In this paper, we have illustrated the substantial bias that can result from failing to do so.

ACKNOWLEDGMENTS

Author affiliations: Department of Biostatistics, University of Washington School of Public Health, Seattle, Washington (Kevin C. Cain); Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, Michigan (Siobán D. Harlow, Matheos Yosef); Department of Biostatistics, University of Michigan School of Public Health, Ann Arbor, Michigan (Roderick J. Little, Bin Nan, Michael R. Elliott); and Centre for Developmental Psychiatry and Psychology, Monash University, Clayton, Victoria, Australia (John R. Taffe).

This work was supported by grant AG021543 from the National Institute of Aging.

Conflict of interest: none declared.

REFERENCES

1. Johnson ES, Matanoski GM. SMR estimations in ‘‘prevalent’’ cohorts and ’’incident‘‘ cohorts. *Med Lav*. 1987;78(4): 263–271.
2. Klein JP, Moeschberger ML. *Survival Analysis: Techniques for Censored and Truncated Data*. New York, NY: Springer; 1997.
3. Kalbfleisch JD, Prentice RL. *The Statistical Analysis of Failure Time Data*. Hoboken, NJ: Wiley; 2002.
4. Meister R, Schaefer C. Statistical methods for estimating the probability of spontaneous abortion in observational studies—analyzing pregnancies exposed to coumarin derivatives. *Reprod Toxicol*. 2008;26(1):31–35.
5. Howards PP, Hertz-Picciotto I, Poole C. Conditions for bias from differential left truncation. *Am J Epidemiol*. 2007; 165(4):444–452.
6. Harlow SD, Cain K, Crawford S, et al. Evaluation of four proposed bleeding criteria for the onset of late menopausal transition. *J Clin Endocrinol Metab*. 2006;91(9):3432–3438.
7. Treloar AE, Boynton RE, Behn BG, et al. Variation of the human menstrual cycle through reproductive life. *Int J Fertil*. 1967;12(1):77–126.
8. Mitchell ES, Woods NF, Mariella A. Three stages of the menopausal transition from the Seattle Midlife Women's Health Study: toward a more precise definition. *Menopause*. 2000;7(5):334–349.
9. Taffe JR, Dennerstein L. Menstrual patterns leading to the final menstrual period. *Menopause*. 2002;9(1):32–40.
10. Soules MR, Sherman S, Parrott E, et al. Executive summary: Stages of Reproductive Aging Workshop (STRAW). *Fertil Steril*. 2001;76(5):874–878.
11. Harlow SD, Crawford S, Dennerstein L, et al. Recommendations from a multi-study evaluation of proposed criteria for staging reproductive aging. *Climacteric*. 2007;10(2): 112–119.
12. Jiang H, Fine JP, Chappell R. Semiparametric analysis of survival data with left truncation and dependent right censoring. *Biometrics*. 2005;61(2):567–575.
13. Taffe JR, Cain KC, Mitchell ES, et al. "Persistence" improves the 60-day amenorrhea marker of entry to late-stage menopausal transition for women aged 40 to 44 years. *Menopause*. 2010;17(1):191–193.
14. Efron B, Tibshirani RJ. *An Introduction to the Bootstrap*. New York, NY: Chapman & Hall; 1993.